magnetic memories, thin-film transistors, low-noise electron tubes, and super-conductive parametric and storage devices. In 1965 he performed an experiment that first proved the existence of the Magnus force in superconductors. In 1966 he became Director of Advanced Memory Devices at Electronic Memories, Inc., Hawthorne, CA, heading the development of plated wire memories. In 1969 he joined the school of Engineering and Applied Science, University of California, Los Angeles, where he is a Professor of Engineering. His present interests lie in signal processing, pattern recognition, artificial intelligence, and decision theory. He has published over 40 technical papers in his fields of interest.

Dr. Pearl is a member of the Association for Computing Machinery (SIGART). In 1965 he was a corecipient of the RCA Research Award for the development of a superconductive parametric amplifier.

A Sinusoidal Family of Unitary Transforms

ANIL K. JAIN, MEMBER, IEEE

***Abstract*—A new family of unitary transforms is introduced. It is shown that the well-known discrete Fourier, cosine, sine, and the Karhunen-Loeve (KL) (for first-order stationary Markov processes) transforms are members of this family. All the member transforms of this family are sinusoidal sequences that are asymptotically equivalent. For finite-length data, these transforms provide different approximations to the KL transform of the said data. From the theory of these transforms some well-known facts about orthogonal transforms are easily explained and some widely misunderstood concepts are brought to light. For example, the near-optimal behavior of the even discrete cosine transform to the KL transform of first-order Markov processes is explained and, at the same time, it is shown that this transform is not always such a good (or near-optimal) approximation to the above-mentioned KL transform. It is also shown that each member of the sinusoidal family is the KL transform of a unique, first-order, non-stationary (in general), Markov process. Asymptotic equivalence and other interesting properties of these transforms can be studied by analyzing the underlying Markov processes.**

***Index Terms*—Fast transforms, image transforms, Karhunen-Loeve transform, orthogonal transforms, unitary matrices.**

I. Introduction

SINCE the discovery of fast Fourier transform (FFT) algorithms [1], [2], interest in unitary and orthogonal transforms, which yield similar fast computational algorithms, has increased rapidly. Many different transforms, e.g., Walsh-Hadamard, Haar, slant, cosine, and sine [3], [4], have been discovered.[1] To a certain extent, the discovery of many unitary transforms has been triggered by attempts to find fast, FFT-type or other, computational algorithms for the family of Karhunen-Loeve (KL) transforms. Since a KL transform is constructed from the eigenvectors of a covariance matrix (of the data that are to be transformed), there is no single unique KL transform for all random processes. Much of the effort in studying KL transforms has been directed at the KL transform of first-order stationary Markov processes whose covariance matrix is given by the Toeplitz matrix R defined as

$$R_{i,j} = \rho^{|i-j|}, \quad |\rho| < 1 \tag{1}$$

where $1 \leqslant i, j \leqslant N$. Although the eigenvectors of R are known analytically [5], they are expensive to generate computationally and have no known fast algorithm to transform a vector of data. Intuitively, the reason is that these eigenvectors, being of the sinusoidal form

$$\phi_m(k) = a_m \sin(w_m k + \theta_m), \quad 1 \leqslant k, \ m \leqslant N \tag{2}$$

where $\{w_m\}$ = solutions of a transcendental equation, are nonharmonic (i.e., the frequency w_m is a nonlinear function of m), which prohibits the existence of an FFT-type algorithm. For higher order Markov processes, closed-form solution of the KL transform is not known in general, and the possibility of a fast algorithm seems even more remote. Recently, Ahmed *et al.* [6] have shown empirically that the discrete cosine transform performs very close to the KL transform of (2) for values of ρ [see (1)] near 0.9. Hamidi and Pearl [7] have extended these results and have shown that the cosine transform yields better performance than the discrete Fourier transform (DFT) for all positive values of ρ. Asymptotic results pertaining to cosine and KL transforms for all finite-order Markov processes have been discussed by Yemini and Pearl [8]. These and other experimental results have led to a widely extrapolated belief that the cosine transform is the best substitute (among the currently known fast unitary transforms) for the KL transform for all first-order, stationary Markov processes. The results of this paper will show that such is not the case. In fact, we shall see that even for first-order Markov processes, there is not unique fast transform within the class examined that will always yield the best approximation to its KL transform. For a higher order Markov

Manuscript received December 30, 1977; revised December 15, 1978. This work was supported in part by the Naval Ocean Systems Center, San Diego, CA, under Contract N00953-77-C-0033MJE and in part by Army Research Office Grant DAAG29-78-G-0206.

The author is with the Department of Electrical Engineering, University of California, Davis, CA 95616.

[1]For a unified treatment of these transforms, see Fino and Algazi [22].

0162-8828/79/1000-0356$00.75 © 1979 IEEE

TABLE I
SOME MEMBERS OF THE SINUSOIDAL TRANSFORM FAMILY

No.	J MATRIX PARAMETERS	TRANSFORM	EIGENVECTORS $\phi_m(k)$, $1 \le m,k \le N$	EIGENVALUES λ_m	Δ	Δ_c COMMUTING DISTANCE
1	$k_1=\rho=k_2$ $k_3=k_4=0$	KLT for 1^{st} order stationary Markov Process	$a_m \sin(\omega_m k+\theta_m)$	Solution of a Transcendental Equation	0	0
2	$k_1=k_2=0$ $k_3=k_4=-1$	DFT	$\frac{1}{\sqrt{N}} \exp \pm [\frac{i2\pi(m-1)(k-1)}{N}]$	$1-2\alpha \cos\frac{2\pi(m-1)}{N}$	$2\alpha^2(1+\rho^2)$	$8\alpha^4(1+\rho^2)$
3	$k_1=k_2=0$ $k_3=k_4=0$	EVEN SINE -1 (EDST-1)	$\sqrt{\frac{2}{N+1}} \sin \frac{mk\pi}{N+1}$	$1-2\alpha \cos\frac{m\pi}{N+1}$	$2\rho^2\alpha^2$	$4\rho^2\alpha^4$
4	$k_1=k_2=1$ $k_3=k_4=0$	EVEN COSINE -1 (EDCT-1)	$\frac{1}{\sqrt{N}}$ $m=1$, $1 \le k \le N$ $\sqrt{\frac{2}{N}} \cos \frac{(2k-1)(m-1)\pi}{2N}$, $m = 2, \ldots N$	$1-2\alpha \cos\frac{(m-1)\pi}{N}$	$2(1-\rho)^2\alpha^2$	$4(1-\rho)^2\alpha^4$
5	$k_1=k_2=-1$ $k_3=k_4=0$	EVEN SINE -2 (EDST-2)	$\sqrt{\frac{2}{N}} \sin \frac{(2k-1)m\pi}{2N}$, $m \ne N$ $\frac{1}{\sqrt{N}}$, $m = N$	$1-2\alpha \cos\frac{m\pi}{N}$	$2(1+\rho)^2\alpha^2$	$4(1+\rho)^2\alpha^4$
6	$k_1=0$, $k_2=+1$ $k_3=k_4=0$	ODD SINE -1 (ODST-1)	$\frac{2}{\sqrt{2N+1}} \sin \frac{(2m-1)k\pi}{2N+1}$	$1-2\alpha \cos\frac{(2m-1)\pi}{2N+1}$	$(1-\rho)^2\alpha^2+\rho^2\alpha^2$	$2(1-\rho)^2\alpha^4+2\rho^2\alpha^4$
7	$k_1=0$, $k_2=-1$ $k_3=k_4=0$	ODD SINE -2 (ODST-2)	$\frac{2}{\sqrt{2N+1}} \sin \frac{2mk\pi}{2N+1}$	$1-2\alpha \cos\frac{2m\pi}{2N+1}$	$(1+\rho)^2\alpha^2+\rho^2\alpha^2$	$2(1+\rho)^2\alpha^4+2\rho^2\alpha^4$
8	$k_1=+1$, $k_2=0$ $k_3=k_4=0$	ODD COSINE -1 (ODCT-1)	$\frac{2}{\sqrt{2N+1}} \cos \frac{(2k-1)(2m-1)\pi}{2(2N+1)}$	$1-2\alpha \cos\frac{2m\pi}{2N+1}$	$(1-\rho)^2\alpha^2+\rho^2\alpha^2$	$2(1-\rho)^2\alpha^4+2\rho^2\alpha^4$
9	$k_1=-1$, $k_1=0$ $k_3=k_4=0$	ODD SINE -3 (ODST-3)	$\frac{2}{\sqrt{2N+1}} \sin \frac{(2k-1)m\pi}{2N+1}$	$1-2\alpha \cos\frac{2m\pi}{2N+1}$	$(1+\rho)^2\alpha^2+\alpha^2\rho^2$	$2(1+\rho)^2\alpha^4+2\rho^2\alpha^4$
10	$k_1=+1$, $k_2=-1$ $k_3=k_4=0$	EVEN COSINE -2 (EDCT-2)	$\sqrt{\frac{2}{N}} \cos \frac{(2k-1)(2m-1)\pi}{4N}$	$1-2\alpha \cos\frac{(2m-1)\pi}{2N}$	$2\alpha^2(1+\rho^2)$	$4\alpha^4(1+\rho^2)$
11	$k_1=-1$, $k_2=+1$ $k_3=k_4=0$	EVEN SINE -3 (EDST-3)	$\sqrt{\frac{2}{N}} \sin \frac{(2k-1)(2m-1)\pi}{4N}$	$1-2\alpha \cos\frac{(2m-1)\pi}{2N}$	$2\alpha^2(1+\rho^2)$	$4\alpha^4(1+\rho^2)$

process, therefore, one may not even expect to find a single best substitute for its KL transform. Instead, one may find the best substitute for a KL transform only for a certain set of values for the statistical parameters of the underlying random process. Alternatively, one may model the random process in such a way that its KL transform is a fast-transform member of the sinusoidal family.

The sinusoidal transforms introduced here, and specifically the sine transform members (see Table I), are different from the previously introduced "fast KL transform algorithm" [9]-[12], which has also been called the "pinned KL transform" [13]. This algorithm is based on the fact that certain random processes can be decomposed as a sum of two mutually orthogonal processes and can be written as

$$u = u^o + u^b \tag{3}$$

where u^b, called the boundary response, is completely determined by the boundary values of the random process sample functions, and the KL transform of the residual process $u - u^b = u^o$ is a sine transform. This sine transform has a fast computational algorithm associated with it and is a member of the sinusoidal family introduced here. This sine transform is not the KL transform of the original random process u, but is the KL transform of the modified process u^o. Rate distortion calculations of the fast KL transform algorithm for data compression of first-order stationary Markov processes have shown superior performance than the cosine transform. However, this algorithm and related considerations are not the subject of this paper. Details may be found in [9]-[12] and [23].

II. THE SINUSOIDAL TRANSFORM FAMILY

Consider the parametric family of matrices

$$J = J(k_1, k_2, k_3, k_4) = \begin{bmatrix} 1 - k_1\alpha & -\alpha & & k_3\alpha \\ -\alpha & 1 & \ddots & \\ & \ddots & 1 & -\alpha \\ k_4\alpha & & -\alpha & 1 - k_2\alpha \end{bmatrix}. \tag{4}$$

This is a variation of the well-known tridiagonal Jacobi matrix.[2] For $k_3 = k_4$, J is a symmetric matrix, and for suitably

[2] A matrix J is called a Jacobi matrix if $J_{m,n} = 0$, for $(m - n) \geqslant 2$ [14]. Hence, for $k_3 = k_4 = 0$, (4) is a Jacobi matrix.

chosen α it would be admissable as a positive definite covariance matrix. For $k_3 = k_4 = 0$, and

$$k_1 = \rho, \quad |\rho| < 1$$
$$\alpha \triangleq \rho/(1+\rho^2)$$
$$\beta^2 = (1-\rho^2)/(1+\rho^2) \tag{5}$$

it can be shown that

$$J(\rho, \rho, 0, 0) = \beta^2 R^{-1} \tag{6}$$

where R is the covariance matrix of the stationary, first-order Markov process defined in (1). Since the eigenvectors of a matrix are invariant under all commuting transformations, the eigenvectors of R^{-1} [and hence of $J(\rho, \rho, 0, 0)$] and R are identical. Since $J(\rho, \rho, 0, 0)$ is also a covariance matrix, the KL transform of its underlying random process is the same as the KL transform of the stationary first-order Markov process, which can be written as

$$u_{k+1} = \rho u_k + \epsilon_k, \quad k = 1, \cdots, N-1$$
$$E[u_1^2] = 1 \quad E[u_1 \epsilon_1] = 0$$
$$E[\epsilon_k] = 0 \quad E[\epsilon_k \epsilon_l] = (1-\rho^2)\delta_{k,l}. \tag{7}$$

Now consider all those J matrices in (4) that are admissable as covariance matrices, i.e., for those values of $k_1, k_2, k_3 = k_4$, and α for which J is positive definite. *The sinusoidal family of unitary transforms is the class of complete orthonormal sets of eigenvectors generated by these J matrices.* Note that each covariance matrix guarantees an associated complete orthonormal set of eigenvectors. Table I summarizes some of the sinusoidal transforms obtained by solving for orthonormal vector solutions of

$$J\phi_m = \lambda_m \phi_m, \quad 1 \leqslant m \leqslant N \tag{8}$$

for different sets of k_i, $1 \leqslant i \leqslant 4$. In this table ϕ_m represents the mth column of the sinusoidal transform matrix Φ_m. All of the eleven transforms listed in Table I are different in the sense that they are eigenvectors of different noncommuting matrices. It is clear that the well-known transforms such as the KLT (number 1), the DFT (number 2), the sine transform (number 3) [10], and the cosine transform (number 4) [6] are members of this family. From (4), (8), and the fact that $k_3 = k_4$, all the sinusoidal transforms are solutions of the homogeneous second-order difference equation

$$\phi_m(k) - \alpha[\phi_m(k-1) + \phi_m(k+1)]$$
$$= \lambda_m \phi_m(k), \quad 2 \leqslant k \leqslant N-1 \tag{9a}$$

subject to the parametric family of boundary conditions

$$(1-k_1\alpha)\phi_m(1) - \alpha\phi_m(2) + k_3\alpha\phi_m(N) = \lambda_m \phi_m(1)$$
$$k_3\alpha\phi_m(1) - \alpha\phi_m(N-1) + (1-k_2\alpha)\phi_m(N) = \lambda_m\phi_m(N). \tag{9b}$$

Thus all the sinusoidal transforms satisfy the same difference equation [i.e., 9(a)] and differ only in the boundary conditions (9b). Table I also lists the eigenvalues associated with the different transforms. These are, of course, the eigenvalues of the J matrices. Interestingly, two different sinusoidal transforms may yield identical eigenvalues, e.g., the J matrices corresponding to ODST-2 (number 7) and ODCT-1 (number 8) have identical sets of eigenvalues, although their eigenvectors are different. The fact that sinusoidal transforms are solutions of (9a) and (9b) can be checked by direct substitution. The unitary property of these transforms can also be verified easily by showing that $\Phi\Phi^{*T} = \Phi^{*T}\Phi = I$. The proofs of these results are straightforward and are obtained by assuming the general solution of (9a) to be of the form

$$\phi_m(k) = A \exp\{ik\theta\} + B \exp\{-ik\theta\}$$

where A, B, and θ are complex in general, and depend on the value of m. These are determined by applying boundary conditions of (9b).

III. Properties of the Sinusoidal Transforms

The following properties of the sinusoidal family can either be proved or are evident from Table I.

A. Orthonormal Properties

The vectors $\{\phi_m(k)\}$ of each family member form a complete orthonormal set of basis vectors in an N-dimensional vector space for $1 \leqslant m, k \leqslant N$. If Φ is the matrix whose mth column is the eigenvector $\{\phi_m(k), 1 \leqslant k \leqslant N\}$, then

$$J\Phi = \Phi\Lambda$$

or

$$\Phi^{*T} J \Phi = \Lambda, \quad \Lambda = \text{diagonal}\,\{\lambda_i, 1 \leqslant i \leqslant N\} \tag{10a}$$

where * denotes the complex conjugate, T denotes the matrix transpose, and $\{\lambda_i\}$ are the eigenvalues of J. Hence, if x is an $N \times 1$ vector whose covariance matrix is $f(J)$ for an arbitrary function $f(\cdot)$, then the sinusoidal transformation

$$\hat{x} = \Phi^{*T} x \tag{10b}$$

is its *Karhunen-Loeve* (KL) transformation. Notice that if Φ is not Hermitian, i.e., $\Phi \neq \Phi^{*T}$, then $\hat{x} = \Phi x$ will not be the KL transformation of x. For example, the EDCT-1, EDST-2, ODST-1, ODST-3 are nonsymmetric, even though they are orthogonal. Hence, the transform domain variances of $\hat{x} = \Phi x$ and $\hat{x} = \Phi^T x$ will be different. This can be seen from Table I by comparing the ODST-1 (number 6) and ODST-3 (number 9). These two transform matrices are transposes of each other. Yet their J matrices as well as the eigenvalues $\{\lambda_m\}$ (which represent the variances of $\hat{x}_m$) are different. Thus, if a given transform were a good approximation to a KL transform (of some arbitrary random process), its transpose need not be so.

B. Comparison with KL Transform

Comparison among the sinusoidal family members for various applications may be made by comparing their J matrices. For example, various transforms could be compared with respect to the transform number 1 in Table I, which is the KL transform of a first-order stationary Markov process. Suppose we wish to know how close the various sinusoidal transforms are to this KL transform. Let us define the difference norm

$$\Delta = \| J(k_1, k_2, k_3, k_4) - J(\rho, \rho, 0, 0) \| \tag{11}$$

where $\|A\| = \Sigma_{i,j} a_{i,j}^2$ is defined as the weak norm of a matrix A. This norm is indicative of the distance between the covariance matrix corresponding to a sinusoidal transform and the covariance matrix corresponding to the KLT in question. [Note that $J(\rho, \rho, 0, 0)$ is not the covariance matrix of the above-mentioned Markov process; rather, it is proportional to the inverse of that covariance matrix; see (8).]

Table I lists the expression for Δ for different sinusoidal transforms. From this it is evident that

$$\Delta(\text{EDST-1}) \geqslant \Delta(\text{EDCT-1}) \quad \text{for} \quad 0.5 \leqslant \rho \leqslant 1 \tag{12a}$$

$$\Delta(\text{EDCT-1}) \geqslant \Delta(\text{EDST-1}) \quad \text{for} \quad -0.5 \leqslant \rho \leqslant 0.5 \tag{12b}$$

$$\Delta(\text{EDCT-1}) \geqslant \Delta(\text{EDST-1}) \geqslant \Delta(\text{EDST-2}) \quad \text{for} \quad -1 \leqslant \rho \leqslant -0.5 \tag{12c}$$

$$\Delta(\text{ODCT-1}) \geqslant \Delta(\text{EDCT-1}) \quad \text{for} \quad 0.5 \leqslant \rho \leqslant 1. \tag{12d}$$

Here, (12a) implies the even cosine-1 transform is better than the even sine-1 transform only for $0.5 \leqslant \rho \leqslant 1$; (12b) and (12c) imply the even sine-1 and even sine-2 transforms perform better than even cosine-1 for other values of ρ; (12d) implies the even cosine-1 performs better than the odd cosine-1 for $0.5 \leqslant \rho \leqslant 1$. Similar inequalities for other sinusoidal transforms can be derived quite straightforwardly.

The validity of these inequalities and hence of the difference norm of (11) as a measure of performance of different transforms can be tested by comparing the actual performances of these transforms. One method of comparing different transforms with a KL transform is to compare their data compression ability. If for any unitary transform Φ and a covariance matrix R we define

$$\sigma_k^2 = [\Phi^{*T} R \Phi]_{k,k} \tag{13}$$

then $\{\sigma_k^2\}$ represents the variances of the transform domain elements of $\hat{x} = \Phi^{*T} x$. If the elements of $\hat{x}$ are ranked in a decreasing order of their variances, i.e., $\sigma_1^2 \geqslant \sigma_2^2 \geqslant \cdots \geqslant \sigma_N^2$, then we will call the quantity

$$J_m = \sum_{k=m+1}^{N} \sigma_k^2 \Big/ \sum_{k=1}^{N} \sigma_k^2 \tag{14}$$

as the "basis restriction error" of the transform Φ for a restriction m. It represents the normalized, minimum mean square error of restricting $\hat{x}$ to m degrees of freedom (i.e., $\hat{x}$ is allowed to have only m nonzero elements). Since $\{\sigma_k^2\}$ depend on the transform Φ via (13), the basis restriction error, for any given m, will vary with Φ and will be minimum for the KL transform. In order to test the inequalities of (12a)-(12d), we let R be given by (1). Figs. 1-3 show comparisons of basis restriction errors of the various transforms for different values of ρ. In the case when a DFT is used, the transformed vector $\hat{x}$ exhibits a conjugate symmetry, i.e.,

$$\hat{x}_k = \hat{x}_{N-k+2}^*, \quad k = 2, \cdots, N$$

$$\hat{x}_1 = \hat{x}_1^*, \quad \hat{x}_{(N/2)+1} = \hat{x}_{(N/2)+1}^*.$$

This gives

$$\sigma_k^2 = \sigma_{N-k+2}^2, \quad k = 2, \cdots, N$$

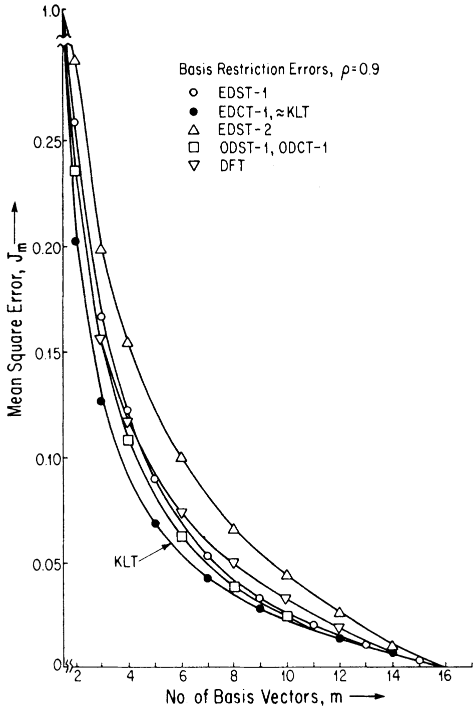


Fig. 1. Basis restriction errors of sinusoidal transforms for a first-order, stationary Markov process of duration $N = 16$, and correlation $\rho = 0.9$.

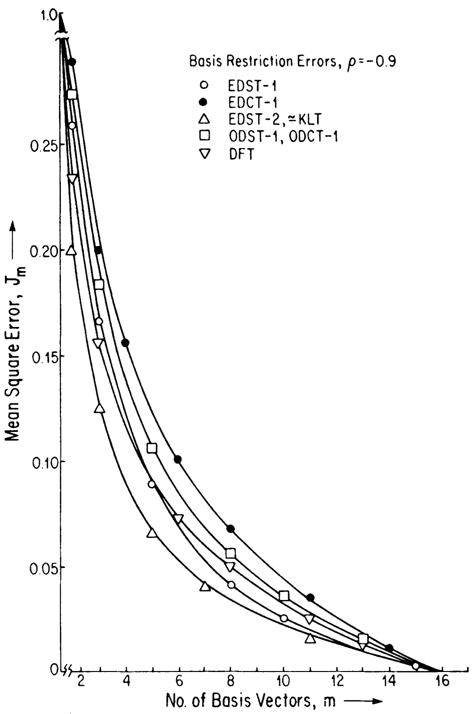


Fig. 2. Basis restriction errors of sinusoidal transforms for a first-order, stationary Markov process of duration $N = 16$, and correlation $\rho = -0.9$.

which results in relatively poor performance of the DFT. An improved ordering of σ_k^2 (with respect to basis restriction error) is obtained by ordering the real and imaginary points of $\hat{x}_k$ according to their variances. This amounts to a permutation between the real and imaginary parts of the vectors of DFT and yields an improved performance. In Figs. 1-4, we have considered this latter ordering.

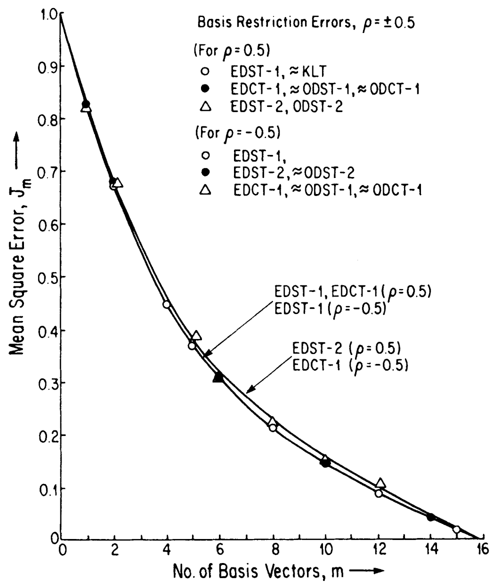


Fig. 3. Basis restriction errors of sinusoidal transforms for a first-order, stationary Markov process of duration $N = 16$, and correlation $\rho = \pm 0.5$.

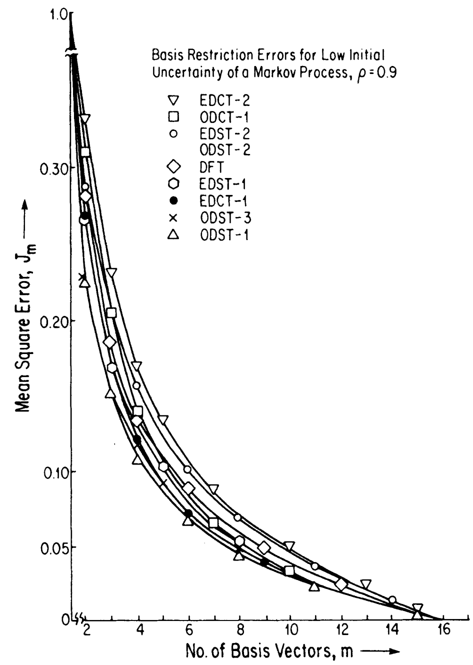


Fig. 4. Basis restriction errors of sinusoidal transforms for a first-order, stationary Markov process of duration $N = 16$, and correlation $\rho = 0.9$, and starting with zero initial condition.

For $\rho = 0.9$, we get the previously reported results of Ahmed *et al.* [6] where the EDCT-1 (generally called, simply, the cosine transform) performs very close to the KLT. This is also expected from our inequality of (12a) and from Table I by noting that Δ(EDCT-1) is minimum (except for the case of the KL transform itself) whenever $0.5 \leqslant \rho < 1$. However, when $\rho = -0.9$, the cosine transform EDCT-1 has the worst performance!

Fig. 4 shows the results for a first-order homogeneous Markov process that starts with zero initial condition and $\rho = 0.9$. Here we find that the odd sine transform (ODST-1) gives the best performance. The various transforms perform in accordance with (12c), and other similar inequalities may be concluded using the criterion of (11). For $|\rho| < 0.5$, the inequality (12b) is also satisfied. However, the performance differences among the various transforms are marginal. These results show that the performance ranking induced by Δ is validated by the performance ranked by the basis restriction error J_m. The inferior performance of the cosine transform, EDCT-1, for $\rho < 0.5$, with respect to some of the other sinusoidal transforms, contradicts the widely held belief that the former is always the best approximation to the KL transform of (2). Although the previous work on cosine transform, such as the empirical results of Ahmed *et al.* [6], who studied its properties for values of $\rho \simeq 0.9$, some asymptotic results and comparisons of cosine transform and DFT reported by Yemini and Pearl [8] and Hamidi and Pearl [7] (where a different criterion is used), and other experimental studies [15], has not specifically compared the cosine transform with all the other transforms for negative or small positive values of ρ; this widely held belief has emerged by their extrapolation. Our results, based on the simple criterion of (11) [and verified by evaluation of (14)], show that the best approximating transform varies with the value of ρ. This is not unreasonable to expect because, after all, the KL transform basis vectors of (2) also vary with the value of ρ.

C. KL Transform Approximation for Arbitrary Random Processes

From the foregoing discussion, one might wonder how the KL transform of an arbitrary random process might be substituted by a (fast) sinusoidal transform. Let A, B be any two covariance matrices. If these matrices commute (i.e., $AB = BA$), then they have an identical set of eigenvectors. Now, define *commuting distance* between any two A and B matrices as

$$\Delta_c = \|AB - BA\| \qquad \Delta_n = \Delta_c/(\|A\| \cdot \|B\|). \tag{15}$$

We conjecture that Δ_c or Δ_n may be used as a measure of proximity of the eigenvectors of A and B. Table I lists the commuting distances between the KL transform covariance matrix R and other J matrices. Comparisons with Δ values show that Δ_c satisfies the same inequalities as does Δ in (12a)–(12d). The difference norm $\Delta = \|R - J\|$ will not be very meaningful for an arbitrary covariance matrix R and the given J matrix. For example, if $R = J^2$, then R and J have the same eigenvectors, but their difference norm Δ is not zero. The above criterion also suggests that for a random process whose covariance function can be measured, if one could reasonably model $R \simeq f(J)$ for a suitably chosen J matrix (e.g., among those J matrices that yield fast transform), then its KL transform would be a (fast) sinusoidal transform. For example, the function f could be an nth-order polynomial. In general, $f(J)$ will not be a Toeplitz matrix so that many nonstationary random processes may be modeled by it. Finally, since all the J matrices are highly sparse (they are tridiagonal whenever $k_3 = k_4 = 0$), only $0(N^2)$ operations are needed in the calculation of Δ_c or Δ_n.

Example: Consider the symmetric, banded Toeplitz matrix

$$H = \{h_{i,j}\}, \qquad 1 \leqslant i,\ j \leqslant N \tag{16}$$

where $h_{i,j} = h_{i-j} = h_{j-i}$ and $h_k = 0$, $|k| > p$. Such matrices arise frequently in the modeling of stationary random processes. For instance, H could represent the covariance function of a pth-order moving average model. Alternatively, H could be the model of a point-spread function (PSF) of an imaging system [18] (note that H represents a finite impulse response (FIR) of a noncausal system). The matrix H has a decomposition

$$H = f(J) - H_b \tag{17}$$

where $f(x)$ is a pth-order polynomial in x and H_b is a matrix of rank at most $2p$. For example, if J is chosen to be the circulant matrix $J(0, 0, -1, -1)$, then H_b is a matrix given as

$$H_b = \begin{bmatrix} 0 & f^T \\ f & 0 \end{bmatrix} \tag{18}$$

where f is a $p \times p$ lower triangular, Toeplitz matrix, defined as

$$\begin{aligned} f_{i,j} &= f_{i-j} = h_{i-j}, \quad && i \geqslant j \\ &= 0, && i < j. \end{aligned} \tag{19}$$

If J is chosen to be $J(1, 1, 0, 0)$ then, following [16], it can be shown that (17) is satisfied when

$$H_b = \begin{bmatrix} \tilde{g} & 0 \\ 0 & \tilde{g} \end{bmatrix} \tag{20}$$

where $\tilde{g}$ is a $(p-1) \times (p-1)$ Hankel matrix whose elements are

$$\tilde{g}_{i,j} = h_{i+j}, \qquad 1 \leqslant i, \ j \leqslant p. \tag{21}$$

Since $h_k = 0$, for $|k| > p$, (21) implies

$$\tilde{g}_{i,j} = h_{i+j}, \qquad 2 \leqslant i + j \leqslant p. \tag{22}$$

Similarly, the H_b matrices, corresponding to other J matrices, could be found, and the best approximation to the KL transform of H is obtained by minimizing the error

$$\Delta_c = \| JH - HJ \| = \| JH_b - H_b J \|$$

or

$$\Delta_n = \| JH_b - H_b J \| / (\| H \| \cdot \| J \|).$$

To extend the previous example to, say, pth-order homogeneous Markov processes, we note that the covariance matrix of such an $N \times 1$ process can be written as [16]

$$R^{-1} = H + H_d \tag{23}$$

where H_d is a sparse matrix of rank at most $2p$. Specifically, H_d contains $p \times p$ matrix blocks in its upper left and lower right corner and is zero elsewhere. Combining (15) and (23) and defining $\tilde{H}_b = H_b - H_d$, we get

$$R = (f(J) - \tilde{H}_b)^{-1} \tag{24}$$

where $\tilde{H}_b$ is of the form

$$\tilde{H}_b = \begin{bmatrix} g_{11} & 0 & g_{12} \\ 0 & 0 & 0 \\ g_{21} & 0 & g_{22} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & 0 \\ 0 & I_p \end{bmatrix} \cdot \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} I_p & 0 & 0 \\ 0 & 0 & I_p \end{bmatrix} \tag{25}$$

where $\{g_{ij}\}$ are $p \times p$ matrix blocks and I_p is a $p \times p$ identity matrix. Now using the $ABCD$ lemma[3] for inversion of $(A - BCD)$, it can be shown that (24) reduces to

$$R = [f(J)]^{-1} + [f(J)]^{-1} \, \Delta R \, [f(J)]^{-1} \tag{26}$$

where ΔR is a sparse matrix of rank at most $2p$. For the case when J is a circulant matrix, the derivation of (26) is given in [17]. The general result of (26) for any of the J matrices follows by proceeding in a manner similar to the development of [17]. The explicit form of ΔR is not relevant to our discussion here and its derivation is left to the reader. Using (26) in (15) for $A = R, B = J$, we get

$$\Delta_c = \| [f(J)]^{-1} \, (J \Delta R - \Delta R J) \, [f(J)]^{-1} \| \tag{27}$$

which can be used for evaluating the approximation of the KLT (for R) by different transforms. Equation (27) can be simplified considerably by observing that J and $f(J)$ are diagonalized by the sinusoidal transform Φ and Δ_c is invariant under a unitary transformation. Hence

$$\Delta_c = \sum_{i,j} \left[\frac{(\lambda_i - \lambda_j)}{f_i f_j} \, \mu_{ij} \right]^2 \tag{28}$$

where

$$f_i = f(\lambda_i), \qquad \mu_{ij} = [\Phi \Delta R \Phi^{*T}]_{i,j} \tag{29}$$

and $\{\lambda_i\}$ are the eigenvalues of J. Since ΔR is a sparse matrix and Φ could be a fast transform, μ_{ij} can be easily computed.

D. Fast Sinusoidal Transforms

A large number of sinusoidals yield fast algorithms for the transformation $\hat{x} = \phi x$ or $\hat{x} = \phi^{*T} x$. Table II shows algorithms for several transforms via the DFT. Hence, these transforms could be implemented via a suitable FFT algorithm in $O(N \log_2 N)$ operations. In general, all the transforms that are harmonic sinusoidal functions will yield such a fast implementation. Thus it follows that all of these transforms will also yield a chirp z-transform (CZT) implementation [18] that can be performed in real time via charge-transfer and surface acoustic-wave devices. These devices allow sampled analog signals at input and output and perform computations in time duration that are proportional to N rather than $N \log N$ [19]. Equation (26) suggests that the inverse of any nearly banded, Toeplitz matrix of the form of (23) may be obtained easily via the fast sinusoidal transforms. The desired inverse R, given by (26), is obtained by calculation of $[f(J)]^{-1}$ and ΔR. Since $f(J)$ is diagonalized by Φ, $[f(J)]^{-1}$ requires only $O(N \log N)$ operations; see, e.g., [17] when Φ is the DFT and [16] when Φ is the sine transform EDST-1. Calculation of ΔR requires at most $O(p^3)$ operations. This is because ΔR depends only on $2p$ independent quantities. When H_d in (23) is zero, ΔR can be computed in only $O(p^2)$ operations (see, e.g., [17]). This is because the $2p$ equations for determining ΔR become Toeplitz in structure and could be solved in $O(p^2)$ operations via a Levinson–Trench algorithm. Once the representation of (26) is obtained, any calculation of the type $x = Ry$ (which is the solution of the nearly Toeplitz equation $R^{-1} x = y$) can be obtained in $O(N \log N) + O(p^2)$ operations

[3] $(A - BCD)^{-1} = A^{-1} - A^{-1} B (C^{-1} - DA^{-1} B)^{-1} DA^{-1}$.

TABLE II

FAST COMPUTATION ALGORITHM FOR DIFFERENT SINUSOIDAL TRANSFORMS VIA DFT, WHICH IS IMPLEMENTED BY A SUITABLE FFT ALGORITHM

Transform	Computation algorithm for $y_m = \sum_{k=1}^{N} x_k \phi_m(k)$, $m = 1, , , ,N$. x_k is assumed zero for unspecified indices
DFT	$y_m = \mathrm{DFT}\,\{x_k\}_N \overset{\Delta}{=} \frac{1}{\sqrt{N}} \sum_{k=1}^{N} x_k \exp\{\frac{i2\pi(k-1)(m-1)}{N}\}$
EDST-1	$y_m = 2\mathrm{Im}[\mathrm{DFT}\,\{z_k\}_{2N+2}]$, $z_k = x_{k-1}$, $2 \le k \le N+1$
EDCT-1	$y_m = 2\mathrm{Re}[\mathrm{DFT}\,\{z_k\}_{2N} \cdot \exp\{\frac{i(m-1)\pi}{2N}\}]$; $z_k = x_k$, $1 \le k \le N$; $2 \le m \le N$ $y_1 = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} x_k$
EDST-2	$y_{m-1} = 2\mathrm{Im}[\mathrm{DFT}\,\{z_k\}_{2N} \cdot \exp\{\frac{i(m-1)\pi}{2N}\}]$; $z_k = x_k$, $1 \le k \le N$; $2 \le m \le N$ $y_N = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} x_k$
ODST-1	$y_m = 2\mathrm{Im}[\mathrm{DFT}\,\{z_k\}_{2N+1} \cdot \exp\{\frac{i2\pi(m-1)}{2N+1}\}]$; $z_k = x_k \exp\{\frac{ik\pi}{2N+1}\}$, $1 \le k \le N$
ODST-2	$y_{m-1} = 2\mathrm{Im}[\mathrm{DFT}\,\{z_k\}_{2N+1}]$; $z_k = x_{k+1}$, $1 \le k \le N$, $2 \le m \le N+1$
ODCT-1	$y_m = 2\mathrm{Re}[\mathrm{DFT}\,\{z_k\}_{2N+1} \cdot \exp\{\frac{i(2m-1)\pi}{2(2N+1)}\}]$; $z_k = x_k \exp\{\frac{i(k-1)\pi}{2N+1}\}$, $1 \le k \le N$
ODST-3	$y_{m-1} = 2\mathrm{Im}[\mathrm{DFT}\,\{z_k\}_{2N+1} \cdot \exp\{\frac{i(m-1)\pi}{2N+1}\}]$; $z_k = x_k$, $1 \le k \le N$, $2 \le m \le N+1$
EDCT-2	$y_m = 2\mathrm{Re}[\mathrm{DFT}\,\{z_k\}_{2N} \cdot \exp\{\frac{i(2m-1)\pi}{2N}\}]$; $z_k = x_k \exp\{\frac{i\pi(k-1)}{2N}\}$, $1 \le k \le N$
EDST-3	$y_m = 2\mathrm{Im}[\mathrm{DFT}\,\{z_k\}_{2N} \cdot \exp\{\frac{i(2m-1)\pi}{2N}\}]$; $z_k = x_k \exp\{\frac{i\pi(k-1)}{2N}\}$, $1 \le k \le N$

via the fast algorithm. This follows because $z \triangleq [f(J)]^{-1} y$ requires $0(N \log N)$ operations and ΔRz requires only $0(p^2)$ operations. Thus, these fast transforms could be useful in substituting the KL transform of certain random processes and also in fast and exact solutions of sparse, nearly Toeplitz, matrix inversion problems.

E. Asymptotic Equivalence

All of the sinusoidal transforms are asymptotically equivalent (see Section V). This can be seen intuitively by inspection of the J matrices. Since the J matrices differ only at boundary points (i.e., the corner elements), as the size of these matrices gets larger, the boundary effects get smaller. For (doubly) infinite J matrices, the boundary positions go to infinity and all the J matrices are the same. This equivalence can also be established by showing asymptotic equivalence of random processes underlying the J matrices and is done in the next section. It should be pointed out that asymptotic equivalence does not imply that for small (or even practical) values of N these transforms perform quite closely. For example, Fig. 1 shows that performances at compression of $2 (m = 8)$ for the $N = 16$ case can be different by about 1.6 to 2 dB when $\rho = 0.9$. These differences will be larger for higher values of ρ and will be even larger for two-dimensional signals, such as images. In fact, performance differences for images are large enough to warrant the use of the best approximating fast transform.

IV. THE UNDERLYING MARKOV PROCESSES

In general, each sinusoidal transform is the KL transform of a nonstationary, first-order Markov process. From Section II, we know that the J matrix of (6) is related to the covariance matrix of the stationary Markov process of (7). The KL transform of this process is the (slow) transform defined in (2). Now we wish to identify the Markov processes that correspond to the general J matrix. First, assume $k_3 = k_4 = 0$. Let us define a noncausal representation for a stochastic process x as

$$Jx = \nu \tag{30}$$

where ν is a zero mean random process vector whose covariance is given by

$$R_\nu = E[\nu\nu^T] = \beta^2 J. \tag{31}$$

From (30), it follows easily that

$$R_x = E[xx^T] = \beta^2 J^{-1}. \tag{32}$$

Comparison with (6) shows that $R_x = R$ when $J = J(\rho, \rho, 0, 0)$. Equations (30) and (31) can be written explicitly as

$$x_k - \alpha(x_{k-1} + x_{k+1}) = \nu_k, \quad 2 \leqslant k \leqslant N - 1 \tag{33a}$$

$$(1 - k_1\alpha) x_1 - \alpha x_2 = \nu_1 \tag{33b}$$

$$(1 - k_2\alpha) x_N - \alpha x_{N-1} = \nu_N \tag{33c}$$

and

$$E[\nu_k \nu_l] = \beta^2(\delta_{k,l} - \alpha\delta_{k-1,l} - \alpha\delta_{k+1,l}). \tag{34}$$

The representation of (30) or, equivalently, (33) is a minimum variance representation of a random process $\{x_k\}$ whose covariance is $\beta^2 J^{-1}$. This means that the linear combination

$$\bar{x}_k = \alpha(x_{k-1} + x_{k+1}), \qquad 2 \leqslant k \leqslant N-1$$

$$\bar{x}_1 = \alpha x_2/(1 - k_1\alpha)$$

$$\bar{x}_N = \alpha x_N/(1 - k_2\alpha)$$

constitutes the best linear mean square estimate of x_k obtained from all possible neighbors of x_k. In other words, the mean square value of the error $\nu_k = x_k - \bar{x}_k$; i.e., $E[\nu_k^2]$ is minimized with the above definitions of $\bar{x}_k$. This follows by observing that the orthogonality relation

$$E[\bar{x}_k \nu_k] = 0, \qquad \text{for all } k$$

which implies $E[x_l \nu_k] = 0$, $k \neq l$, and $E[x_k \nu_k] = E[\nu_k^2] = \beta^2$, is satisfied because, from (30) and (31), we get

$$E[x\nu^T] = J^{-1}E[\nu\nu^T] = \beta^2 I.$$

If $\{x_k\}$ were to be a first-order Markov process, it must have a representation

$$x_k = r_k x_{k-1} + s_k, \qquad 2 \leqslant k \leqslant N \tag{35}$$

where $\{r_k\}$ and $\{s_k\}$ are deterministic and random white-noise sequences, respectively. Repeated substitution of (35) in (33a) yields the identity

$$[(1 - \alpha r_{k+1}) r_k - \alpha]\, x_{k-1} - \alpha s_{k+1} + (1 - \alpha r_{k+1}) s_k = \nu_k. \tag{36}$$

Now, we let the sequence $\{r_k\}$ be such that

$$(1 - \alpha r_{k+1}) r_k - \alpha = 0$$

or

$$r_k = (1 - \alpha r_{k+1})^{-1}\alpha, \qquad k = 2, \cdots, N-1. \tag{37}$$

Then (36) reduces to

$$-\alpha s_{k+1} + (1 - \alpha r_{k+1}) s_k = \nu_k, \qquad 2 \leqslant k \leqslant N. \tag{38}$$

For (38) to be a valid realization for the random process $\{\nu_k\}$, (34) must be satisfied. Defining

$$E[s_k s_l] = p_k \delta_{k,l} \tag{39}$$

and substituting (38) in (34), using (39), and solving for p_k, we get

$$p_k = \frac{\beta^2 r_k}{\alpha}, \qquad k = 3, \cdots, N-1. \tag{40}$$

Using (34) in (32c), we obtain

$$r_N = \alpha/(1 - k_2\alpha), \qquad |r_N| < 1 \quad \text{for} \quad |k_2| < 1 \tag{41}$$

and

$$s_N = \nu_N/(1 - k_2\alpha). \tag{42}$$

For (35) to be a Markov representation we must assume x_1 and s_2 to be independent. The variances of the initial state x_1 and random variable s_2 are then obtained from (42) and (34) as

$$E[x_1^2] = \beta^2/(1 - \alpha r_2 - k_1\alpha) \tag{43a}$$

$$E[s_2^2] = (r_2 + k_1)/\alpha. \tag{43b}$$

This completely specified (35). Thus, we see that for each J matrix there is a nonstationary (in general) Markov process given by (35) whose KL transform is determined by the eigenvectors of J. The covariance of the random process x is $\beta^2 J^{-1}$. Thus all the fast sinusoidal transforms are KL transforms of nonstationary Markov processes; whereas, the (slow) transform of (2) is the KL transform of the stationary Markov process of (7).

V. Asymptotic Equivalence

The transforms of the sinusoidal family can be shown to be asymptotically equivalent. To establish this equivalence we first consider some definitions following Gray [20] and Pearl [21].

1) Let A_N be a real $N \times N$ matrix. Then

$$\eta_w(A_N) = \eta_w \triangleq \left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} a_{i,j}^2\right)^{1/2} \tag{44}$$

is called the *weak* (*or Hilbert-Schmidt*) *norm* of A_N. If $\{\lambda_k\}$ are the eigenvalues of $A_N^T A_N$, then

$$\eta_s(A_N) = \max_k \{\lambda_k\} \tag{45}$$

is called the *strong norm* of A_N.

2) Two sequences of matrices $\{A_N\}$ and $\{B_N\}$, $N = 1, 2, \cdots$ are considered asymptotically equivalent, written as $A_N \sim B_N$, if

$$\lim_{N\to\infty} \eta_w(A_N - B_N) = 0. \tag{46}$$

3) The matrix A_N is weakly or strongly bounded if $\eta_w(A_N) < \infty$ or $\eta_s(A_N) < \infty$, respectively. Based on these definitions we can state the following useful theorem [20].

A. Theorem 1

Let $\{A_N\}$ and $\{B_N\}$ be strongly bounded, real, symmetric matrix sequences with

$$\lambda_{\min} \leqslant \eta_s(A_N), \qquad \eta_s(B_N) \leqslant \lambda_{\max}. \tag{47}$$

Then, if $\{A_N\}$ and $\{B_N\}$ are asymptotically equivalent, and for an arbitrary function $F(x)$, continuous on $[\lambda_{\min}, \lambda_{\max}]$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} F(\alpha_{N,k}) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} F(\beta_{N,k}) \tag{48}$$

where $\{\alpha_{N,k}\}$ and $\{\beta_{N,k}\}$ are the eigenvalues of A_N and B_N, respectively. This relation is valid if either of the two limits exists.

This theorem also implies, for any $m < \infty$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=m+1}^{N} F(\alpha_{N,k}) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=m+1}^{N} F(\beta_{N,k}). \tag{49}$$

Thus, if $F(x) = x$, (49) states that the basis restriction errors in data compression of two classes of random processes whose covariances are A_N and B_N are asymptotically equivalent. For $F(x) = \log_2(x/\theta)$, $(0 < \theta < x, \forall x > 0)$ the same conclusion can be drawn from (48) for the rate-distortion functions of these two random processes.

4) Let $\{A_N\}$ and $\{B_N\}$ be a sequence of real strongly bounded convariance matrices with corresponding KL transform unitary matrices $\{\Phi_N\}$ and $\{\psi_N\}$, respectively. Define

$$\hat{\Lambda}_B = \text{diag}\,\{\Phi_N B_N \Phi_N^{*T}\}$$
$$\hat{B}_N = \Phi_N^{*T} \hat{\Lambda}_B \Phi_N.$$

The matrix $\hat{B}_N$ is called the projection of B_N on the sequence $\{A_N\}$. Note that both $\hat{B}_N$ and A_N are diagonalized by Φ_N.

5) The transform sequence $\{\Phi_N\}$ is said to asymptotically cover the sequence $\{A_N\}$ if A_N and $\hat{B}_N$ are asymptotically equivalent, i.e., $A_N \sim \hat{B}_N$.

6) The transform sequences $\{\Phi_N\}$ and $\{\psi_N\}$ are said to be asymptotically equivalent if $A_N \sim \hat{B}_N$ and $B_N \sim \hat{A}_N$.

We are now ready to prove the following asymptotic equivalence results.

a) The J matrices of the sinusoidal family are asymptotically equivalent. This is easily seen by noting that

$$\eta_W^2(\Delta) = \frac{\alpha^2}{N} \{(k_1' - k_1)^2 + (k_2' - k_2)^2 + (k_3' - k_3)^2 + (k_4' - k_4)^2\} \tag{50}$$

where

$$\Delta \triangleq J - J' \triangleq J(k_1, k_2, k_3, k_4) - J(k_1', k_2', k_3', k_4').$$

Clearly, $\eta_W(\Delta) \to 0$ as $N \to \infty$.

b) All the sinusoidal transforms are asymptotically equivalent. Let J_N and J_N' be two J matrices with Φ_N and ψ_N the corresponding transforms.

Define

$$\hat{\Lambda}_N' = \text{diag}\,\{\Phi_N J_N' \Phi_N^{*T}\}.$$

Since

$$J_N' = J_N - \Delta$$

we obtain

$$\Lambda_N' = \Phi_N J_N \Phi_N^{*T} - \text{diag}\,[\Phi_N \Delta \Phi_N^{*T}].$$

This gives

$$\lambda_N'(k) = \lambda_N(k) + \alpha(k_1 - k_1')\,|\phi_N(k,1)|^2 + \alpha(k_2 - k_2') \cdot |\phi_N(k,N)|^2 + \alpha(k_3' - k_3)\,\phi_N(k,1)\,\phi_N^*(k,N) + \alpha(k_4' - k_4)\,\phi_N^*(k,1)\,\phi_N(k,N). \tag{51}$$

From this, one readily obtains

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \lambda_N'(k) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \lambda_N(k)$$

i.e., $\Lambda_N \sim \Lambda_N'$, which implies $J_N \sim \hat{J}_N'$. Similarly, we can show $J_N' \sim \hat{J}_N$. From definition 6), this implies $\{\Phi_N\}$ and $\{\psi_N\}$ are asymptotically equivalent.

c) The Markov realizations of the sinusoidal family are asymptotically equivalent.

In the steady state, i.e., for $N = \infty$, the coefficients r_k satisfy

$$(1 - \alpha r)\, r - \alpha = 0.$$

Letting $\alpha \triangleq \rho/(1 + \rho^2)$, we get

$$r = \rho \quad \text{or} \quad 1/\rho. \tag{52}$$

Since $|\alpha| < \frac{1}{2}$ for a valid covariance matrix J, $r = \rho$, $|\rho| < 1$ is the only admissible solution. Indeed, if $|r_N| < 1$, then from (37) $|r_k| < 1$, $\forall k$, so that r_k converges to ρ. Also, in the steady state

$$E[s_k^2] = p_k = \frac{\beta^2 \rho}{\alpha} = (1 - \rho^2). \tag{53}$$

Hence, (35) becomes the homogeneous Markov process

$$x_k = \rho x_{k-1} + s_k \tag{54}$$

which will be stationary as $k \to \infty$ for arbitrary initial conditions.

For the case of the DFT, $(k_3 = k_4 = -1)$. Letting $\alpha = \rho/(1 + \rho^2)$, we can obtain a circular factorization of $J(0, 0, -1, -1)$ to give a stationary representation

$$x_k = \rho x_{k-1} + s_k, \qquad 2 \leqslant k \leqslant N \tag{55}$$

$$x_1 = \rho x_N + s_1 \tag{56}$$

$$E[s_k s_l] = \beta^2 \delta_{k,l}.$$

As $N \to \infty$, for any arbitrary initial value x_1, this will be equivalent to (54) as $k \to \infty$.

d) The sinusoidal transforms are asymptotically equivalent to the KL transform of any finite-order, stationary Markov process.

The spectral density function of a pth-order, stationary Markov process is given by

$$S(\omega) = \frac{2}{\left(1 - 2 \sum_{k=1}^{p} a_k \cos \omega k\right)}. \tag{57}$$

We will assume $S(\omega)$ is strictly positive. From (23) to (25), the $N \times N$ covariance matrix of a Markov sequence of length N satisfies

$$R_N^{-1} = f(J) + \tilde{H}_b \tag{58}$$

where $\tilde{H}_b$ is a sparse matrix with at most $4p^2$ nonzero terms. Defining

$$R_N^0 = [f(J)]^{-1}$$

it is easy to show that $R_N^{-1} \sim (R_N^0)^{-1}$. The eigenvalues of R_N^0 are given by

$$\lambda_N^0(k) = \beta^2 \Big/ \left(1 - 2 \sum_{l=1}^{p} a_l \cos kl\theta_N\right) \tag{59}$$

where θ_N depends on N and the J matrix.

Defining

$$\hat{D}_N = \operatorname{diag}\,[\Phi R_N^{-1} \Phi^{*T}], \qquad \hat{H}_N = \Phi^{*T} \hat{D}_N \Phi$$

we obtain

$$\hat{D}_N = (\Lambda_N^0)^{-1} + \operatorname{diag}\,(\Phi \tilde{H}_b \Phi^{*T}). \tag{60}$$

Using (25), one obtains $\hat{D}_N \sim (\Lambda_N^0)^{-1}$ or $\hat{H}_N \sim f(J)$. Similarly, we can show that $\hat{f}(J) \sim R_N^{-1}$. This means Φ is asymptotically equivalent to the KL transform of a pth-order Markov process.

V. Conclusions

In conclusion, a family of sinusoidal transforms has been introduced. All the transform family members are generated by a class of J matrices that is a slight variation of the Jacobi matrices. Equivalently, all the sinusoidal transforms are solutions of a second-order eigenvalue difference equation, subject to different boundary conditions. Each sinusoidal transform is the KL transform of a first-order Markov process. The fast sinusoidal transforms are the KL transforms of nonhomogeneous (and hence nonstationary) random processes, whereas, for the stationary Markov process, the corresponding KLT is not fast. Applications of the transform family discussed here are in the modeling of random processes such that their KL transforms are fast transforms and are also in fast solutions of sparse Toeplitz or nearly Toeplitz systems of equations. Both of these applications could be useful in many signal-processing problems, such as data compression and digital filtering of signals. Although all the sinusoidal transforms are asymptotically equivalent, their performance differences for signals with finite support can be significant for certain ranges of signal parameters.

References

[1] J. W. Cooley and J. W. Tukey, "An algorithm for machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, pp. 297–301, 1965.

[2] E. O. Brigham, *The Fast Fourier Transform*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

[3] N. Ahmed and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*. New York: Springer, 1975.

[4] W. K. Pratt, *Image Processing*. Reading, MA: Addison-Wesley, 1977.

[5] W. D. Ray and R. M. Driver, "Further decomposition of the Karhunen-Loeve series representation of a stationary random process," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 663–668, Nov. 1970.

[6] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. C-23, pp. 90–93, Jan. 1974.

[7] M. Hamidi and J. Pearl, "Comparison of the cosine and Fourier transforms of Markov-1 signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 428–429, Oct. 1976.

[8] Y. Yemini and J. Pearl, "Asymptotic properties of discrete unitary transforms," School of Eng. and Appl. Sci., Univ. of Calif., Los Angeles, UCLA-ENG-Rep. 7566, Nov. 1975.

[9] A. K. Jain, "Computer program for fast Karhunen Loeve transform algorithm," Dep. Elec. Eng., State Univ. of New York, Buffalo, Final Rep. NASA Contract NAS8-31434, Feb. 1976.

[10] ——, "A fast Karhunen Loeve transform for a class of stochastic processes," *IEEE Trans. Commun.*, vol. COM-24, pp. 1023–1029, Sept. 1976.

[11] A. K. Jain, S. H. Wang, and Y. Z. Liao, "Fast Karhunen Loeve transform data compression studies," in *Proc. National Telecommunications Conf.*, Dallas, TX, Nov. 1976.

[12] A. K. Jain, "Some new techniques in image processing," in *Image Science Mathematics*, C. O. Wilde and E. Barrett, Eds. North Hollywood: Western Periodicals, Nov. 1976, pp. 201–223.

[13] A. Z. Meiri, "The pinned Karhunen-Loeve transform of a two-dimensional Gauss Markov field," in *Proc. 1976 SPIE Conf. Image Processing*, San Diego, CA, Aug. 1976.

[14] R. Bellman, *Introduction to Matrix Analysis*. New York: McGraw-Hill, 1970.

[15] W. C. Chen and C. H. Smith, "Adaptive coding of color images using cosine transform," in *Proc. Int. Communications Conf.*, June 1976, pp. 47-7–47-13.

[16] A. K. Jain, "An operator factorization method for restoration of blurred images," *IEEE Trans. Comput.*, vol. C-25, pp. 1061–1071, Nov. 1977.

[17] ——, "Fast inversion of banded Toeplitz matrices by circular decompositions," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 121–127, 1978.

[18] L. R. Rabiner *et al.*, "The chirp *Z* transform algorithm," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, pp. 86–92, 1969.

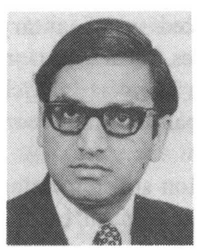
[19] H. J. Whitehouse, R. W. Means, and J. M. Speiser, "Signal processing using transversal filter technology," Naval Undersea Center, San Diego, CA, Tech. Rep., 1975.

[20] R. M. Gray, "Toeplitz and circulant matrices: A review," Stanford Univ., Stanford, CA, Tech. Rep. SU-SEL-71-032, June 1971; also ——, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 725–30, Nov. 1972.

[21] J. Pearl, "Asymptotic equivaluence of spectral representations," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-23, pp. 547–551, Dec. 1975.

[22] B. J. Fino and R. Algazi, "A unified treatment of fast unitary transforms," *SIAM J. Comput.*, vol. 6, Dec. 1977.

[23] A. K. Jain, *Multidimensional Techniques in Digital Image Processing*, to be published.



Anil K. Jain (M'74) was born in India on January 21, 1946. He received the B. Tech. (Honors) degree from IIT, Kharagpur, India, in 1967, and the M.S. and Ph.D. degrees from the University of Rochester, Rochester, NY, in 1969 and 1970, respectively, all in electrical engineering.

From 1968 to 1970 he was a System Science Fellow at the University of Rochester, and from 1970 to 1971 he was a Post-Doctoral Fellow with the Departments of Electrical Engineering and Computer Science, University of Southern California, Los Angeles. From 1971 to 1974 he was an Assistant Professor with the Department of Electrical Engineering and the Image Processing Institute, University of Southern California. From September 1974 to 1978 he was an Associate Professor with the Department of Electrical Engineering, State University of New York, Buffalo. Since 1978 he has been an Associate Professor with the Department of Electrical and Computer Engineering, University of California, Davis. He has been a Consultant for several private and federal organizations. His research interests include digital signal and image processing, real time computer systems, and pattern recognition.